

REMARKS

By the present Amendment, the specification has been amended to expressly recite material that was incorporated by reference in the captioned application at the time of filing. Support for the amendments to the specification are found for example at column 13, lines 52-60 of U.S. Patent No. 5,608,144; at column 15, lines 42-51 of U.S. Patent No. 4,563,517; at page 64, lines 9-15 of U.S. Application No. 60/111,990 and U.S. Application No. 09/459,109; and at page 63, lines 9-15 of U.S. Application No. 09/459,110, all of which documents were incorporated by reference in their entireties at the time of filing. No new matter enters by these amendments.

Claims 2, 12, 14, 19-21, 26 and 60-64 are currently amended and new claims 65-78 are added. Claims 24 and 25 have been cancelled without prejudice or disclaimer of the underlying subject matter. Applicants note that their previous Amendment and Response to Office Action filed August 21, 2008 and their previous Supplemental Amendment and Response filed August 26, 2008 included new claim 64, which Applicants respectfully note was not addressed in the Office Action mailed October 3. Claims 1, 3-5, 7-11, 15-18, 22-23 and 27-59 remain cancelled. Claims 6 and 13 are original. Support for the present amendments may be found, for example, in the original claims, sequence listing and specification, for example at page 2, line 22-24 and page 4, lines 2-4. No new matter enters by these amendments.

I. Request for Continued Examination

Applicants thank the Office for the acknowledgment that “the finality of the previous Office action has been withdrawn” and “Applicant’s submission filed on 8/26/2008 has been entered.” Office Action at page 2.

II. Election of Species

Applicants note the Office's comments regarding Applicants' election with traverse of plant cells in a Response to Restriction Requirement filed April 20, 2007. Claims 19-21, 25-26, and 62-64 have been amended to recite "transformed plant cell." *See e.g.*, Claim 19.

III. Rejection under 35 U.S.C. § 112, 2nd paragraph

Claims 2, 6, 12-14, 19-21, 24-26, and 60-63 were rejected under 35 U.S.C. § 112 as allegedly being indefinite. Office Action at page 3. The Office argues that "[i]t is vague and indefinite what is meant by the phrase 'comprising from about 30 to 300. The phrase typically indicates a minimum point. The phrase, however, is contraverted by the term 'about' which implies that values above and below 30 nucleotides are permitted." Office Action at pages 3-4. Applicants respectfully disagree with the Office's argument.

The legal standard for indefiniteness under 35 U.S.C § 112, second paragraph, is whether a claim reasonably apprises those of skill in the art of its scope. *See, Amgen Inc. v. Chugai Pharm. Co.*, 927 F.2d 1200, 1217, 18 USPQ2d 1016, 1030 (Fed. Cir. 1991). In *Amgen*, the court "caution[ed] that our holding that the term "about" renders indefinite claims 4 and 6 should not be understood as ruling out any and all uses of this term in patent claims. It may be acceptable in appropriate fact situations, *e.g.*, *W.L. Gore & Assocs., Inc. v. Garlock, Inc.*, 721 F.2d 1540, 1557, 220 USPQ 303, 316 (Fed.Cir.1983) ("use of 'stretching ... at a rate exceeding about 10% per second' in the claims is not indefinite")." (emphasis added) *Amgen Inc. v. Chugai Pharm. Co.*, 927 F.2d 1200 at 1217.

“The definiteness inquiry focuses on whether those skilled in the art would understand the scope of the claim when the claim is read in light of the rest of the specification.” *Union Pacific Resources Co. v. Chesapeake Energy Corp.*, 236 F.3d 684, 692 (Fed. Cir. 2001). In contrast to the facts in *Amgen*, where the referenced measurement was imprecise, Applicants’ specification provides for a precision and specificity adequate to guide one of skill in the art regarding the meaning of the phrase “at least about” as used to modify a given number of nucleotide residues. As such, this phrase would be readily understood by the skilled artisan, and the rejection under 35 USC 112, 2nd Paragraph, Indefiniteness, should be withdrawn.

IV. Rejections Under 35 U.S.C. § 101

Claims 2, 6, 12-14, 19-21, 24-26 and 60-61 were rejected under 35 U.S.C. § 101 as allegedly “not supported by either a credible asserted utility or a well established utility.” Office Action at page 4. Applicants respectfully disagree with this allegation.

Claim 2 recites, *inter alia*, “[a] substantially purified nucleic acid molecule... comprising at least about 100 nucleotide residues of the nucleic acid sequence of SEQ ID NO: 5272....” Although the Office has acknowledged that according to the specification “...the claimed nucleic acids can be used to determine transcriptional profiling...” and “...can be use for mapping studies, linkage analysis, constructing transgenic plants, and screening for traits or screening for polymorphisms...”, the Office argues that “[t]he instant specification does not contemplate the specific use of SEQ ID No. 5272....” *See e.g.*, Office Action mailed March 18, 2008 at page 7 and Office Action at page 11. The Office further argues that “...at the time of filing ... there had been no contemplation of a specific useful trait for SEQ ID No. 5272.” Office Action at page 11.

Applicants respectfully disagree with the Office's arguments. Applicants' specification recites that identified sequences, which necessarily include for example SEQ ID NO: 5272, can act as regulatory elements and as genes. *See e.g.*, page 1, lines 19-26. Applicants' specification also notes the use of identified sequences to alter yield. *See e.g.*, page 2, lines 17-21. Applicants respectfully point out that the Office must accept these stated utilities in the absence of evidence or sound scientific reasoning to rebut Applicants' assertion. *In re Oetiker*, 977 F.2d 1443, 1445 (Fed. Cir. 1992). Moreover, the present application has been awarded priority to U.S. Provisional Application 60/155,422, filed September 23, 1999 ("the '422 application"), which was incorporated by reference in its entirety at the time of filing the present application. This '422 priority application identifies SEQ ID NO: 5272 (which is referred to as SEQ ID NO: 9911 in the '422 application) as a COL2 gene. *See e.g.*, Attachment D.¹

Furthermore, as of the September 23, 1999 priority date of the captioned application, those skilled in the art were well aware that COL2 referred to a "CONSTANS-like" gene and showed significant homology to CONSTANS, where CONSTANS had been identified as a putative zinc finger transcription factor affecting growth, namely, flowering. *See e.g.*, Putterill, J. *et al.*, Cell 80:847-857 (1995) and Ledger, S.E. *et al.*, PGR 96-081 112:862 (1996). As such, it is respectfully submitted that Applicants had established a specific, substantial, and credible utility for SEQ ID NO: 5272 at the time of filing.

The Office asserts that "SEQ ID No. 9911 is listed as being 34% identical to COL2.

¹ Attachment D contains information excerpted from priority application 60/155,422. For the convenience of the Office, the column headers that appear at the beginning of the table from which Attachment D was excerpted have been carried over from the first page of the table where they appear and added to this Attachment.

Neither the instant specification nor the provisional specification provides any information about which nucleotides are similar between SEQ ID No. 5272 and COL2.” Office Action at pages 13-14. The Office also argues that neither the captioned application nor the provisional priority application “provide any information ... rather (sic.) SEQ ID No. 5272 and COL2 share similar nucleotides at the putative zinc finger transcription site.” Office Action at page 14. However, as shown by the attached alignment, G1988² and COL2 share greater than 44% homology at the protein level within the zinc finger domain. *See* Attachment F.

In the meantime, since filing their application, Applicants have also provided additional evidence further demonstrating that specific and substantial utilities of SEQ ID NO: 5272, as recited at the time of filing the ‘422 priority application, are indeed accurate. For example, U.S. Patent Publication 2008/0010703 evidences the fact that the specific and substantial utilities stated in Applicants’ specification as filed and in the ‘422 priority application are indeed utilities of SEQ ID NO: 5272. *See e.g.*, US 2008/0010703.

As detailed in U.S. Patent Publication 2008/0010703, G1988 is a nucleic acid sequence that differs by a single nucleotide from the corresponding region of SEQ ID NO: 5272. *See e.g.*, Attachment E. However, this nucleotide difference does not alter the encoded protein (*i.e.*, is a silent nucleotide change). As such, G1988 encodes the identical protein as the corresponding region of SEQ ID NO: 5272. *See id.* G1988 has been demonstrated to increase yield in plants in field trials conducted in 3 different years. *See e.g.*, US 2008/0010703 at Figure 6 and Tables 12

² As detailed in U.S. Patent Publication 2008/0010703, G1988 is a nucleic acid sequence that differs by a single nucleotide from the corresponding region of SEQ ID NO: 5272. This single nucleotide difference does not alter the

and 13. Indeed, co-pending U.S. Application No. 11/821,448 evidences “significantly increased yield....” US 2008/0010703 A1 at paragraph [0037].

The Office argues that “though SEQ ID NO. 5272 might have a use after filing by the further experimentation of G1988; this use was not specifically contemplated in the instant specification.” Office Action at page 14. Applicants disagree. As noted previously, the ‘422 priority application identifies SEQ ID NO: 5272 (which is referred to as SEQ ID NO: 9911 in the ‘422 application) as a COL2 gene, and those skilled in the art at the time of filing were well aware that COL2 referred to a “CONSTANS-like” gene, which showed significant homology to CONSTANS, where CONSTANS had been identified as a putative zinc finger transcription factor affecting growth, namely, flowering. Applicants provided a specific and substantial utility for SEQ ID NO: 5272 at the time of filing.

Furthermore, not only did Applicants provide a specific and substantial utility for SEQ ID NO: 5272 at the time of filing as previously discussed, but also in the meantime Applicants have provided evidence that the utilities they provided were indeed accurate. Accordingly, Applicants respectfully note that reference to U.S. Patent Publication 2008/0010703 provides post-filing evidence to support utilities properly asserted by Applicants at the time of filing.

On the basis of the foregoing, Applicants respectfully submit that the claimed nucleotide sequence has utilities specific to it, and not simply general utilities applicable to any nucleic acid. These specific and substantial utilities were provided by Applicants at the time of filing and

encoded protein (*i.e.*, is a silent nucleotide change). As such, G1988 encodes the identical protein as the corresponding region of SEQ ID NO: 5272.

satisfy the requirements of 35 U.S.C. § 101. Moreover, in the meantime, US 2008/0010703 has provided additional evidence demonstrating the accuracy of Applicants' previously asserted utilities.

Based on the foregoing, Applicants respectfully submit that the present application fulfills the legal requirements of 35 U.S.C. § 101, Utility. As such, Applicants request withdrawal of the utility rejection.

V. Rejection under 35 U.S.C. § 112, Enablement

Claims 2, 6, 12-14, 19-21, 24-26 and 60-61 were rejected under 35 U.S.C. § 112, first paragraph, as not being enabled by the specification, because the claimed invention allegedly lacks utility (*i.e.*, an invention with no utility cannot be enabled). Applicants respectfully traverse this rejection, and note that this rejection has been overcome by the foregoing arguments regarding utility. As such, reconsideration and withdrawal of the enablement rejection under 35 U.S.C. § 112, first paragraph is respectfully requested.

VI. Rejections Under 35 U.S.C. § 102, Novelty

a. 102(e) Moyer *et al.* U.S. Publication No. 2005/0014263

Claims 2 and 60-61 were rejected under 35 U.S.C. § 102(e) as allegedly being anticipated by Moyer *et al.* (US Patent Application Publication US2005/0014263 1/20/2005 filing date of 5/29/1998)." Office Action at page 16. Applicants respectfully traverse this rejection.

According to the Office "the phrase 'about 30 to 200 nucleotide residues...' can be interpreted broadly to encompass any sequence which has at least 30 nucleotides in common...."

Office Action at page 16. Applicants respectfully draw the attention of the Office to Claims 2 and 60-61, which have been amended to recite for example, “comprising at least about 100 contiguous nucleotide residues....” *See e.g.*, Claim 2. Accordingly, the Office’s rejection of Claims 2 and 60-61 under 35 U.S.C. § 102(e) over Moyer *et al.* has been rendered moot, and withdrawal of this rejection is respectfully requested.

b. 102(b) GenBank Accession No. N37270

Claims 2 and 60-61 were rejected under 35 U.S.C. § 102(b) as allegedly being anticipated by GenBank Accession No. N37270. Applicants respectfully draw the attention of the Office to claims 2 and 60-61, which recite for example, “comprising at least about 100 contiguous nucleotide residues....” *See e.g.*, Claim 2. Accordingly, Applicants respectfully submit that the presently pending claims cannot be anticipated by N37270, and Applicants request withdrawal of the Office’s rejection under 35 U.S.C § 102(b).

CONCLUSION

In view of the foregoing amendments and remarks, Applicants respectfully submit that the present application is now in condition for allowance, and respectfully request notice of such. The Examiner is encouraged to contact the undersigned at 202-942-5325 if any additional information is necessary for allowance.

Respectfully submitted,

Date: April 2, 2009



David R. Marsh (Reg. No. 41,408)

Lisa A. Adelson (Reg. No. 51,204)

Arnold & Porter LLP
555 Twelfth Street, N.W.
Attn: IP Docketing
Washington, DC 20004

Telephone: 202-942-5000
Facsimile: 202-942-5999

ATTACHMENT D

Seq_Num	Contig Id	Gene Id	Position	Hit Id	AAT nap Score	Blast Score	Blast pvalue	%Ident	%Cvrg	Hit Description
9902	ATL8C11158	ATL8On11837	291-1	g4586265	159	150	1.3e-10	37	38	(AL049640) putative protein [Arabidopsis thaliana]
9903	ATL8C6899	ATL8On11838	1203-686	g4406785	873	872	3.0e-87	94	51	(AC006532) hypothetical protein [Arabidopsis thaliana]
9904	ATL8C11157	ATL8On11839	1-406	g465975	214	286	1.1e-24	52	21	PUTATIVE ATP-DEPENDENT RNA HELICASE T26G10.1 IN CHROMOSOME III [
9905	ATL8C6893	ATL8On11840	1-1761	g3219271	409	557	7.2e-54	43	80	(AB015315) MAP kinase kinase 4 [Arabidopsis thaliana]
9906	ATL8C28545	ATL8On11841	678-522	g3047086	192	227	1.2e-17	71	6	(AF058914) similar to reverse transcriptase (Pfam: transcript_fact.hmm, score: 72.31) [Arabidopsis thaliana]
9907	ATL8C28548	ATL8On11842	762-1	g4741194	458	624	5.7e-61	77	22	(AL049746) ABC transporter-like protein [Arabidopsis thaliana]
9908	ATL8C28549	ATL8On11843	888-408	g4204269	704	505	4.7e-47	81	12	(AC005223) 64111 [Arabidopsis thaliana]
9909	ATL8C28547	ATL8On11844	306-1000	g4417310	457	344	3.2e-31	65	33	(AC006446) putative Athila retroelement ORF1 protein [Arabidopsis thaliana]
9909	ATL8C28547	ATL8On11845	816-1000	g3319352	258	184	4.3e-13	84	7	(AF077407) No definition line found [Arabidopsis thaliana]
9910	ATL8C28544	ATL8On11846	109-2289	g3242708	3325	2860	6.5e-298	97	95	(AC003040) putative serine/threonine protein kinase [Arabidopsis thaliana]
9911	ATL8C44589	ATL8On11847	3172-1136	g1507699	83	161	2.5e-11	34	27	(L81119) COL2 [Arabidopsis thaliana]
9912	ATL8C44586	ATL8On11848	468-225	g2245058	349	361	4.2e-33	90	85	(Z97342) hypothetical protein [Arabidopsis thaliana]
9913	ATL8C44588	ATL8On11849	964-1	g4262158	125	220	1.1e-16	66	5	(AC005275) hypothetical protein [Arabidopsis thaliana]
9914	ATL8C44587	ATL8On11850	1-658	g3421384	727	797	2.7e-79	94	38	(AF081067) IAA-Ala hydrolase; IAA-amino acid hydrolase [Arabidopsis thaliana]
9915	ATL8C11164	ATL8On11851	568-1698	g3510254	1338	675	8.1e-129	69	100	(AC005310) putative zinc transporter [Arabidopsis thaliana]
9915	ATL8C11164	ATL8On11852	1972-2165	g4836911	126	177	9.6e-13	56	15	(AC007153) 55420 [Arabidopsis thaliana]
9916	ATL8C11162	ATL8On11853	162-672	g4263705	377	391	2.8e-36	47	71	(AC006223) putative disease resistance protein [Arabidopsis thaliana]
9917	ATL8C11165	ATL8On11854	2025-1	g3608135	932	773	9.3e-77	75	91	(AC005314) putative DNA binding factor [Arabidopsis thaliana]
9918	ATL8C11167	ATL8On11855	393-653	g2244958	468	468	1.9e-44	100	99	(Z97340) hypothetical protein [Arabidopsis thaliana]
9919	ATL8C11166	ATL8On11856	2925-1267	g3193321	1982	2068	1.6e-219	82	100	(AF069299) No definition line found [Arabidopsis thaliana]
9920	ATL8C11169	ATL8On11857	1213-2993	g4567296	2152	1788	2.6e-184	80	100	(AC006918) putative Athila retroelement ORF1 protein [Arabidopsis thaliana]
9921	ATL8C28550	ATL8On11858	520-1	g4335720	596	655	3.0e-63	74	14	(AC006248) putative reverse transcriptase Tal-1 [Arabidopsis thaliana]
9922	ATL8C11168	ATL8On11859	1-405	g4544460	454	345	2.1e-31	65	42	(AC006592) putative reverse transcriptase [Arabidopsis thaliana]
9923	ATL8C28551	ATL8On11860	729-199	g4467359	882	641	6.8e-62	97	16	(AJ002685) Phosphatidylinositol 4-kinase [Arabidopsis thaliana]
9924	ATL8C28553	ATL8On11861	491-1	g2653885	491	507	7.2e-48	96	13	(AF027408) phospholipase D-gamma; PLD-gamma [Arabidopsis thaliana]
9925	ATL8C28554	ATL8On11862	1742-1	g2864623	1746	1751	2.1e-180	69	36	(AL021811) putative protein [Arabidopsis thaliana]
9925	ATL8C28554	ATL8On11863	547-2215	g2864620	228	271	1.5e-23	72	35	(AL021811) hypothetical protein [Arabidopsis thaliana]
9925	ATL8C28554	ATL8On11864	2521-3748	g2864621	137	169	6.3e-12	48	21	(AL021811) hypothetical protein [Arabidopsis thaliana]
9925	ATL8C28554	ATL8On11865	1332-3748	g4544372	351	454	2.1e-41	42	13	(AC006920) putative reverse transcriptase [Arabidopsis thaliana]
9925	ATL8C28554	ATL8On11866	2658-3748	g4406792	119	223	2.5e-17	39	14	(AC006304) putative reverse transcriptase [Arabidopsis thaliana]
9926	ATL8C44591	ATL8On11867	1824-3724	g2583131	1164	1345	2.3e-137	52	100	(AC002387) putative pectinesterase [Arabidopsis thaliana]
9927	ATL8C28555	ATL8On11868	859-1	g4914332	1090	1094	9.0e-111	77	45	(AC005489) F14N23.18 [Arabidopsis thaliana]

ATTACHMENT E

>nucleotides 2536 to 3210 of SEQ ID NO: 5272

AACGTTGTGCTTTTCAGCCCATCCTTCCTCCGAATCCACGCGCCACCGTCTAAGCTGCTGCGTCATTGCACGCGCCA
ATTTGCTTTCAACCGCTCGAATCATCCAGCTGAACTCCAGTCACATCTTCTACTTTCTTTAAATTTCTGCCACGTC
GTCGTGTTCTTAACGCGCAACCAAAACGCGCGCTAAGAACAATCTCTTCGTGCTCTTGGCCTCGTCTCCACAGC
CAAAGCCAAAGACGCATATGAAACGACAGCGTTTGTAAATCCCTGTTTAGTCCTAACTTACCACACCAATTTACAA
AAATGCCATCCGCCACCGTAACCGCAACGGCCTTGGCATTCACTCTGTTTTCCCTCCCTCGCGCTCTGTTTACGTCA
CGCGTCGTTGACGATAGCTCGGAGCTTGAGACACAGTCAAGAGACGAGCAGCAAGAAGAAGACGACGATTCTGAACA
ACATGTTGTTGCTGGAGGCCAAGGAAGAAGAGGACCAGAAACGAAATTTGAGTAAGAGATTGCAATTTGGGCAA
TGACACGCCGGAATGACGAGCGAAGAGAAAATTTGAGGCATGGAATTTAGCGTCACAAGAACGGCAGAGGAAGGCA
GAGTCCGCGGCACAATGGAGATCAGCTTCGGCACCACAAAGCTCGCAAAGCTCACCAT

>Reverse complement of nucleotides 2536 to 3210 of SEQ ID NO: 5272

ATGGTGAGCTTTTGCGAGCTTTGTGGTGCCGAAGCTGATCTCCATTGTGCCGCGGACTCT
GCCTTCCTCTGCCGTTCTTGTGACGCTAAGTTCCATGCCTCAAATTTTCTCTTCGCTCGT
CATTTCCGGCGTGTCAATTTGCCCAAATTGCAAATCTCTTACTCAAATTTCTGTTTCTGGT
CCTCTTCTTCTTGGCCTCCACGAACAACATGTTGTTTCAAGATCGTCGTCTTCTTCTTGC
TGCTCGTCTCTTGAAGCTGTGTCTCAAGCTCCGAGCTATCGTCAACGACGCGTGACGTAAAC
AGAGCGCGAGGGAGGGAAAACAGAGTGAATGCCAAGGCCGTTGCGGTTACGGTGGCGGAT
GGCATTTTTGTAAATTTGGTGTGGTAAGTTAGGACTAAACAGGGATTTAAACAAACGCTGTC
GTTTCATATGCGTCTTTGGCTTTGGCTGTGGAGACGAGGCCAAGAGCGACGAAGAGAGTG
TTCTTAGCGCGCGCTTTTGGTTCGGCGTTAAGAACACGACGACGTGGCAGAATTTAAAG
AAAGTAGAAGATGTGACTGGAGTTTCACTGGGATGATTGAGCGGTTGAAAGCAAATTTG
CGCGTGCATGACGCAGCAGCTTAGACGGTGGCGCGTGGATTGCGAGGAAGGATGGGCT
GAAAACGACAACGTT

>G1988 in US 2008/0010703 (hereinafter "Publ")

ATGGTGAGCTTTTGCGAGCTTTGTGGTGCCGAAGCTGATCTCCATTGTGCCGCGGACTCTGCCTTCCTCTGCCGTTT
TTGTGACGCTAAGTTCCATGCCTCAAATTTTCTCTTCGCTCGTCATTTCCGGCGTGTCTATCTGCCCAAATTTGCAAAT
CTCTTACTCAAATTTCTGTTTCTGGTCTCTTCTTCTTGGCCTCCACGAACAACATGTTGTTTCAAGATCGTCGTCT
TCTTCTTGTGCTCGTCTCTTGAAGCTGTGTCTCAAGCTCCGAGCTATCGTCAACGACGCGTGACGTAAACAGAGCGCG
AGGGAGGGAAAACAGAGTGAATGCCAAGGCCGTTGCGGTTACGGTGGCGGATGGCATTTTTGTAAATTTGGTGTGGTA
AGTTAGGACTAAACAGGGATTTAAACAAACGCTGTGTTTCATATGCGTCTTTGGCTTTGGCTGTGGAGACGAGGCCA
AGAGCGACGAAGAGAGTGTTCTTAGCGCGCGCTTTTGGTTCGGCGTTAAGAACACGACGACGTGGCAGAATTTAAA
GAAAGTAGAAGATGTGACTGGAGTTTCACTGGGATGATTGAGCGGTTGAAAGCAAATTTGGCGCGTGCAATGACGC
AGCAGCTTAGACGGTGGCGCGTGGATTGCGAGGAAGGATGGGCTGAAAACGACAACGTTTGA

Reverse complement of nucleotides 2536 to 3210 of SEQ ID NO: 5272

ATGGTGAGCTTTTGCGAGCTTTGTGGTGCCGAAGCTGATCTCCATTGTGCCGCGGACTCTGCCTTCCTCT 70

G1988inPubl

ATGGTGAGCTTTTGCGAGCTTTGTGGTGCCGAAGCTGATCTCCATTGTGCCGCGGACTCTGCCTTCCTCT 70

Consensus

ATGGTGAGCTTTTGCGAGCTTTGTGGTGCCGAAGCTGATCTCCATTGTGCCGCGGACTCTGCCTTCCTCT 70

Reverse complement of nucleotides 2536 to 3210 of SEQ ID NO: 5272

GCCGTTCTTGTGACGCTAAGTTCCATGCCTCAAATTTTCTCTTCGCTCGTCATTTCCGGCGTGTCTATCTG 140

G1988inPubl

GCCGTTCTTGTGACGCTAAGTTCCATGCCTCAAATTTTCTCTTCGCTCGTCATTTCCGGCGTGTCTATCTG 140

Consensus

GCCGTTCTTGTGACGCTAAGTTCCATGCCTCAAATTTTCTCTTCGCTCGTCATTTCCGGCGTGTCTATCTG 140

Reverse complement of nucleotides 2536 to 3210 of SEQ ID NO: 5272

CCCAAATTTGCAAATCTCTTACTCAAATTTTCTGTTTCTGGTCTCTTCTTCTTGGCCTCCACGAACAACA 210

G1988inPubl
CCCAAATTGCAATCTCTTACTCAAAATTTCTGTTTCTGGTCCTCTTCTTCCTTGGCCTCCACGAACAACA 210

Consensus
CCCAAATTGCAATCTCTTACTCAAAATTTCTGTTTCTGGTCCTCTTCTTCCTTGGCCTCCACGAACAACA 210

Reverse complement of nucleotides 2536 to 3210 of SEQ ID NO: 5272
TGTTGTTTCAGAATCGTCGTCTTCTTCTTGCTGCTCGTCTCTTGACTGTGTCTCAAGCTCCGAGCTATCGT 280
G1988inPubl
TGTTGTTTCAGAATCGTCGTCTTCTTCTTGCTGCTCGTCTCTTGACTGTGTCTCAAGCTCCGAGCTATCGT 280

Consensus
TGTTGTTTCAGAATCGTCGTCTTCTTCTTGCTGCTCGTCTCTTGACTGTGTCTCAAGCTCCGAGCTATCGT 280

Reverse complement of nucleotides 2536 to 3210 of SEQ ID NO: 5272
CAACGACGCGTGACGTAACAGAGCGCGAGGGAGGGAAAAACAGAGTGAATGCCAAGGCCGTTGCGGTTAC 350
G1988inPubl
CAACGACGCGTGACGTAACAGAGCGCGAGGGAGGGAAAAACAGAGTGAATGCCAAGGCCGTTGCGGTTAC 350

Consensus
CAACGACGCGTGACGTAACAGAGCGCGAGGGAGGGAAAAACAGAGTGAATGCCAAGGCCGTTGCGGTTAC 350

Reverse complement of nucleotides 2536 to 3210 of SEQ ID NO: 5272
GGTGGCGGATGGCATTTCATTTGTAATTTGGTGTGGTAAGTTAGGACTAAACAGGGATTAAACAAACGCTGTC 420
G1988inPubl
GGTGGCGGATGGCATTTCATTTGTAATTTGGTGTGGTAAGTTAGGACTAAACAGGGATTAAACAAACGCTGTC 420

Consensus
GGTGGCGGATGGCATTTCATTTGTAATTTGGTGTGGTAAGTTAGGACTAAACAGGGATTAAACAAACGCTGTC 420

Reverse complement of nucleotides 2536 to 3210 of SEQ ID NO: 5272
GTTTCATATGCGTCTTTGGCTTTGGCTGTGGAGACGAGGCCAAGAGCGACGAAGAGAGTGTTCCTTAGCGG 490
G1988inPubl
GTTTCATATGCGTCTTTGGCTTTGGCTGTGGAGACGAGGCCAAGAGCGACGAAGAGAGTGTTCCTTAGCGG 490

Consensus
GTTTCATATGCGTCTTTGGCTTTGGCTGTGGAGACGAGGCCAAGAGCGACGAAGAGAGTGTTCCTTAGCGG 490

Reverse complement of nucleotides 2536 to 3210 of SEQ ID NO: 5272
CGGCGTTTTGGTTTCGGCGTTAAGAACACGACGACGCTGGCAGAATTTAAAGAAAGTAGAAGATGTGACTGG 560
G1988inPubl
CGGCGTTTTGGTTTCGGCGTTAAGAACACGACGACGCTGGCAGAATTTAAAGAAAGTAGAAGATGTGACTGG 560

Consensus
CGGCGTTTTGGTTTCGGCGTTAAGAACACGACGACGCTGGCAGAATTTAAAGAAAGTAGAAGATGTGACTGG 560

Reverse complement of nucleotides 2536 to 3210 of SEQ ID NO: 5272
AGTTTCAGCTGGGATGATTTCGAGCGGTTGAAAGCAAATTGGCGCGTGCAATGACGCAGCAGCTTAGACGG 630
G1988inPubl
AGTTTCAGCTGGGATGATTTCGAGCGGTTGAAAGCAAATTGGCGCGTGCAATGACGCAGCAGCTTAGACGG 630

Consensus

AGTTTCAGCTGGGATGATTGAGCGGTTGAAAGCAAATTGGCGCGTGCAATGACGCAGCAGCTTAGACGG 630

Reverse complement of nucleotides 2536 to 3210 of SEQ ID NO: 5272

TGGCGCGTGGATTGCGAGGAAGGATGGGCTGAAAACGACAACGTT--- 675

G1988inPubl

TGGCGCGTGGATTGCGAGGAAGGATGGGCTGAAAACGACAACGTTTGA 678

Consensus

TGGCGCGTGGATTGCGAGGAAGGATGGGCTGAAAACGACAACGTTtga 678

> Protein Sequence encoded by Reverse complement of nucleotides 2536 to 3210 of SEQ ID NO: 5272 (hereinafter "Prot Seq 5272 RC")

MVSFCELCGAEADLHCAADSAFLCRSCDAKFHASNFLFARHFRRVICPNCKSLTQNFVSG
 PLLPWPPTTCCSESSSSSSCCSSSLDCVSSSELSSSTTRDVNRARGREN RVNAKAVAVTVAD
 GIFVNWCCKLGLNRDLTNAVVSASYASLALAVETRPRATKRVFLAAAFWFGVKNTTTWQNLK
 KVEDVTGVSAGMIRAVESKLARAMTQQLRRWRVDSEEGWAENDNV

>Protein Sequence G1988 in Publ

MVSFCELCGAEADLHCAADSAFLCRSCDAKFHASNFLFARHFRRVICPNCKSLTQNFVSG
 PLLPWPPTTCCSESSSSSSCCSSSLDCVSSSELSSSTTRDVNRARGREN RVNAKAVAVTVAD
 GIFVNWCCKLGLNRDLTNAVVSASYASLALAVETRPRATKRVFLAAAFWFGVKNTTTWQNLK
 KVEDVTGVSAGMIRAVESKLARAMTQQLRRWRVDSEEGWAENDNV*

Prot Seq 5272 RC

MVSFCELCGAEADLHCAADSAFLCRSCDAKFHASNFLFARHFRRVICPNCKSLTQNFVSGPLLPWPPTT 70

Protein Sequence G1988 in Publ

MVSFCELCGAEADLHCAADSAFLCRSCDAKFHASNFLFARHFRRVICPNCKSLTQNFVSGPLLPWPPTT 70

Consensus

MVSFCELCGAEADLHCAADSAFLCRSCDAKFHASNFLFARHFRRVICPNCKSLTQNFVSGPLLPWPPTT 70

Prot Seq 5272 RC

CCSESSSSSSCCSSSLDCVSSSELSSSTTRDVNRARGREN RVNAKAVAVTVADGIFVNWCCKLGLNRDLTNAV 140

Protein Sequence G1988 in Publ

CCSESSSSSSCCSSSLDCVSSSELSSSTTRDVNRARGREN RVNAKAVAVTVADGIFVNWCCKLGLNRDLTNAV 140

Consensus

CCSESSSSSSCCSSSLDCVSSSELSSSTTRDVNRARGREN RVNAKAVAVTVADGIFVNWCCKLGLNRDLTNAV 140

Prot Seq 5272 RC

VSYASLALAVETRPRATKRVFLAAAFWFGVKNTTTWQNLKKVEDVTGVSAGMIRAVESKLARAMTQQLRR 210

Protein Sequence G1988 in Publ

VSYASLALAVETRPRATKRVFLAAAFWFGVKNTTTWQNLKKVEDVTGVSAGMIRAVESKLARAMTQQLRR 210

Consensus

VSYASLALAVETRPRATKRVFLAAAFWFGVKNTTTWQNLKKVEDVTGVSAGMIRAVESKLARAMTQQLRR 210

Prot Seq 5272 RC

WRVDSEEGWAENDNV 225

Protein Sequence G1988 in Publ
WRVDSEEGWAENDNV 225

Consensus
WRVDSEEGWAENDNV 225

ATTACHMENT F

COL2	MLKEESNESGTWARA	DT	PSAACTVYCEADSA	TT	DARVHAANRVASRRHERVR	QS	EEAPAAFL	70
G1988	-----MVSF	EL	CAADLCADSA	ET	DALCHASNTTAAHTRRV	DK	LTQNFV	58

Consensus	70
-----------	-------	-------	-------	-------	-------	-------	-------	----

COL2	CKADAASLCTACDAEIHSA	NPARRHQ	VPILPL	AN	CS	MAPSETDADNDEDD	EWASWLLPNPKN	139
G1988	S-----	G	PLPWP	PTCCSE	SSS	CC	SLDCVSSSELSSTT	DN
							---RARE	106

Consensus	140
-----------	-------	-------	-------	-------	-------	-------	-------	-----

COL2	IGNQNGFLFG	HEYLDL	DYSSMDNQFEDNQYTHQ	SFGGDG	VPLOVEESTSHLQSQNQFQ	GINY	209
G1988	NRVNAKAVAVT	ADGIF	NWCGKIG	-----LN	DLTNAV	SYASIALAVETRPRATKRVF	AAAF
							166

Consensus	210
-----------	-------	-------	-------	-------	-------	-------	-------	-----

COL2	GFSSGAHYNNNSL	DLNHSAS	MSMDISVPEST	SDI	IV	QHP	TTKETIDQLSGPPTQVVQQLTPMERE	279
G1988	WFGVKNTTTWQN	KVEDVTG	AGM	RA	ESKL	RAM	QILR	
							WRVDSEEGWAENDNV	
							-----	225

Consensus	280
-----------	-------	-------	-------	-------	-------	-------	-------	-----

COL2	ARVLRREKKTRKFDK	TIRYASRKAYAEIR	PRIKGRFAKRIETEA	EAEEIF	STSLMSETGYGIVPSF	347
G1988	-----	-----	-----	-----	-----	225

Consensus	348
-----------	-------	-------	-------	-------	-------	-------	-------	-----

Matrix of pairwise sequence identities

Color codes for percent identity:	≥0	≥43	≥60	≥74	≥85	≥93
-----------------------------------	----	-----	-----	-----	-----	-----

#	Sequence	1	2
1 COL2	-		14.1 (49)
2 G1988		21.8 (49)	-

Alignment created by *abc_Aligner* (A. Boukharov, Monsanto Bioinformatics, 2005)